数理解析研究所講究録
981 巻 1997 年 22-35

Global and Polynomial-Time Convergence of an Infeasible-Interior-Point Algorithm Using Inexact Computation *

Shinji Mizuno † Florian Jarre ‡

April, 1996

Abstract

In this paper, we propose an infeasible-interior-point algorithm for solving a primal-dual linear programming problem. The algorithm uses inexact computations for solving a linear system of equations at each iteration. Under a very mild assumption on the inexactness we show that the algorithm finds an approximate solution of the linear program or detects infeasibility of the program. The assumption on the inexact computation is satisfied if the approximation to the solution of the linear system is just a little bit "better" than the trivial approximation 0. We also give a sufficient condition to achieve polynomial-time convergence of the algorithm.

Key Words: Linear Programming Problem, Interior-Point Method, Primal-Dual, Convergence, Inexact Computation

1 Introduction

Since the announcement of the projective scaling algorithm by Karmarkar [2], interior-point algorithms have developed tremendously. Most work per iteration of an interior-point algorithm is devoted to the computation of a search direction, which is a solution of a linear system of equations. When the linear system is very large, the evaluation of the solution by a direct method typically requires a lot of computer time. In such a case, one may wish to compute only an approximate solution by using an iterative method. Even if one uses a direct method, the solution may not satisfy the linear equations exactly, because of computational errors. In spite of the inexactness of the solution in practical computations, most analyses of interior-point algorithms have been done under the assumption that we do compute the exact solution of the linear system.

In this paper, we only assume that we compute an approximate solution of the linear system. Our assumption on the approximate solution is so general that it is satisfied if the approximation is just a little bit "better" than the trivial approximation 0. Under this assumption, we propose a primal-dual interior-point algorithm which can start from an infeasible interior point, and we prove its global linear convergence. This type of algorithm is called an infeasible-interior-point algorithm. It was proposed by Lustig et al. [4] and Tanabe [7], and its convergence was proved by Kojima et al. [3] when exact computations are used. We also give a sufficient condition for the inexact computation, under which the number of iterations of

*Research Memorandum 605, The Institute of Statistical Mathematics, Tokyo, Japan

†The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106, Japan. This author's research was supported in part by Grant-in-Aid for Scientific Research (C-07640343) and the ISM Cooperative Research Program (95-ISM.CRP-B1)

‡Institut für Angewandte Mathematik und Statisitk, Universität Würzburg, 97074 Würzburg, Germany

our algorithm is bounded by a polynomial function. The complexity is compatible with the bound of the infeasible-interior-point algorithms proposed by Zhang [8] and Mizuno [5].

This work was stimulated by a related paper [1] in which also inexact search directions are investigated. The analyses in both papers, however, are completely different. The algorithm in [1] is based on a numerical implementation. The algorithm of this paper is of more theoretical nature and therefore allows to prove a somewhat stronger convergence result under a slightly weaker assumption on the inexact computations.

In Section 2, we introduce the linear system to be solved at each iteration of an interior-point algorithm, and we make our assumption on the inexact computations. In Section 3, we present our algorithm using inexact computations and state our main result. In Section 4, we discuss the main result and introduce some new notation. In Section 5, we prove global convergence of our algorithm. In Section 6, we give a sufficient condition to achieve polynomial time convergence.

2 Inexact Computations

We consider a linear programming problem

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b, \quad x \geq 0, \end{array}$$

where A is an $m \times n$ matrix, $b \in R^m$, $c \in R^n$, and $x \in R^n$. We assume that the rank of A is m. This problem is called primal. The dual of the primal problem is defined by

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & A^T y + s = c, \quad s \geq 0, \end{array}$$

where $y \in R^m$ and $s \in R^n$ are variables. We define the primal-dual linear programming problem, which is to find a solution of the system

$$\begin{pmatrix} Ax \\ A^T y + s \\ Xs \end{pmatrix} = \begin{pmatrix} b \\ c \\ 0 \end{pmatrix}, \; x \geq 0, \; s \geq 0,$$

where $X = \operatorname{diag}(x)$ is the diagonal matrix whose diagonal elements are equal to the elements of x. It is well-known that (x, y, s) is a solution of the primal-dual linear programming problem if and only if x and (y, s) are optimal solutions of the primal and dual linear programming problem respectively. We call (x, y, s) a feasible solution if $Ax = b$, $A^T y + s = c$, and $(x, s) \geq 0$.

We will measure the "size" of the right hand sides b and c relative to A by

$$\|(b,c)\|_A = \left\| \begin{pmatrix} b \\ c \end{pmatrix} \right\|_A := \min_{x,y,s} \left\{ \left\| \begin{pmatrix} x \\ s \end{pmatrix} \right\| \; : \; \begin{array}{rcl} Ax & = & b \\ A^T y + s & = & c \end{array} \right\}.$$

Thus, $\|(b,c)\|_A$ is the Euclidean norm of the "smallest" vector pair (x, s) that satisfies the primal-dual equality constraints while ignoring the nonnegativity constraints. Likewise, we will also measure the norm of certain perturbations $\tilde{b}$ and $\tilde{c}$. The above definition implies that $\| \, . \, \|_A$ is a semi-norm (satisfies all norm-properties except that $\|z\|_A = 0$ does not imply $z = 0$), and $\left\|(\tilde{b}, \tilde{c})\right\|_A = \left\|(\tilde{b}, \; A^T y + \tilde{c})\right\|_A$ for any three vectors $\tilde{b}, \tilde{c}$ and y.

A primal-dual interior-point algorithm generates a sequence of points $(x^k, y^k, s^k) \in R^{n+m+n}$ for $k = 0, 1, \cdots$. At the k-th iteration of the algorithm, we solve a system of linear equations

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S_k & 0 & X_k \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta s \end{pmatrix} = \begin{pmatrix} p \\ q \\ r \end{pmatrix}, \tag{1}$$

where $X_k = \operatorname{diag}(x^k)$ and $S_k = \operatorname{diag}(s^k)$ for the current iterate (x^k, y^k, s^k), and (p, q, r) is a vector in R^{n+m+n}.

Suppose that we have an approximate solution $(\Delta x'', \Delta y'', \Delta s'')$ of System (1). Then we can get another approximate solution $(\Delta x', \Delta y', \Delta s')$, which exactly satisfies the third equality $S_k\Delta x' + X_k\Delta s' = r$ as follows:

$$\begin{array}{rcl} \Delta x_i' & = & \begin{cases} \Delta x_i'' + (r_i - s_i^k\Delta x_i'' - x_i^k\Delta s_i'')/s_i^k & \text{if} \quad s_i^k \geq x_i^k, \\ \Delta x_i'' & \text{otherwise,} \end{cases} \\ \Delta y' & = & \Delta y'', \\ \Delta s_i' & = & \begin{cases} \Delta s_i'' & \text{if} \quad s_i^k \geq x_i^k, \\ \Delta s_i'' + (r_i - s_i^k\Delta x_i'' - x_i^k\Delta s_i'')/x_i^k & \text{otherwise.} \end{cases} \end{array} \tag{2}$$

If we set $(\Delta x'', \Delta y'', \Delta s'') = (0,0,0)$ in this procedure, we see that

$$\left\| \begin{pmatrix} A\Delta x' - p \\ A^T\Delta y' + \Delta s' - q \end{pmatrix} \right\|_{\mathcal{A}} = \left\| \begin{pmatrix} A\Delta x' - p \\ \Delta s' - q \end{pmatrix} \right\|_{\mathcal{A}} \leq \left\| \begin{pmatrix} p \\ q \end{pmatrix} \right\|_{\mathcal{A}} + \|U_k^{-1}r\|,$$

where $U_k = \operatorname{diag}(u^k)$ and u^k is defined by

$$u_i^k = \max\{x_i^k, s_i^k\} \text{ for each } i.$$

We will show in Lemma 9 below that $\|U_k^{-1}\|$ is bounded throughout our algorithm if the primal-dual linear programming problem is feasible. Usually one can compute a much better approximate solution than $(0,0,0)$, but we only use the following weak assumption in this paper.

Assumption 1 *We can compute an approximate solution* $(\Delta x', \Delta y', \Delta s')$ *of System (1) such that*

$$\left\|(A\Delta x' - p,\ \Delta s' - q)\right\|_{\mathcal{A}} \leq \sigma_1 \|(p,q)\|_{\mathcal{A}} + \sigma_2\|r\|,$$

and

$$S_k\Delta x' + X_k\Delta s' = r,$$

where $\sigma_1 \in [0,1)$ *and* $\sigma_2 \geq 0$ *are constants independent of* (x^k, y^k, s^k).

If $\sigma_1 = 0$ and $\sigma_2 = 0$, this assumption implies that we can compute the exact solution of the system (1). We do not assume that σ_1 or σ_2 are small except for Section 6.

Assumption 1 on the inaccuracy of the search direction is substantially more general than the inaccuracy which is introduced for example by the rank-1-update proposed in Karmarkar's original paper [2]. In particular, Assumption 1 does not preserve the feasibility of the iterates even if the initial point happens to be strictly primal and dual feasible.

3 A Conceptual Algorithm

In this section we define an interior-point algorithm. We call this algorithm a conceptual algorithm since we do not specify exactly how to compute some of the quantities that are needed in the algorithm. This issue is addressed in part in the next section.

Let (x^0, y^0, s^0) be an initial point such that

$$\rho_0 e \leq x^0 \quad \text{and} \quad \rho_0 e \leq s^0,$$

where $\rho_0 > 0$ is a constant. We call (x^0, y^0, s^0) an (infeasible) interior point for the primal-dual problem since it strictly satisfies the inequality constraints $x \geq 0$ and $s \geq 0$. We define

$$\begin{array}{rcl} \mu^0 & = & (x^0)^T s^0/n, \\ \bar{b} & = & Ax^0 - b, \\ \bar{c} & = & A^T y^0 + s^0 - c. \end{array}$$

For each $\theta > 0$, we consider a system of equations and inequalities

$$\begin{pmatrix} Ax \\ A^Ty+s \\ Xs \end{pmatrix} = \begin{pmatrix} b+\theta\bar{b} \\ c+\theta\bar{c} \\ \theta\mu^0 e \end{pmatrix}, \; x>0,\; s>0. \tag{3}$$

If the primal-dual linear programming problem is feasible, then the system (3) has a unique solution for each $\theta \in (0,1]$, otherwise $\theta_\ell \in (0,1)$ exists such that the system has a unique solution for each $\theta \in (\theta_\ell, 1]$ and does not have a solution for each $\theta \le \theta_\ell$, see Mizuno et al. [6] for example. We call the solution of the system (3) a center, and define the set of centers in the space combining (x,y,s) and θ:

$$\begin{aligned} \mathcal{P} \;=\; & \{(x,y,s,\theta) : x>0,\; s>0,\; \theta>0, \\ & Ax = b+\theta\bar{b},\;\; A^Ty+s = c+\theta\bar{c},\;\; Xs=\theta\mu^0 e\}. \end{aligned}$$

It is well-known that the set $\mathcal{P}$ forms a path, which is called a path of (infeasible) centers, and that (x,y,s) converges to a solution of the primal-dual linear programming problem if $(x,y,s,\theta)\in\mathcal{P}$ and $\theta\to 0$.

Let γ_0, γ_1, γ_2, γ_3, and γ_4 be positive constants such that

$$\gamma_0 < 1 < \gamma_1,\quad \gamma_3<1,\quad \gamma_4<1,\quad \sigma_1\gamma_3+\sigma_2\gamma_2<\gamma_3,$$

$$\gamma_0\mu^0 e \le X_0 s^0 \le \gamma_1\mu^0 e \;\text{ and }\; \|X_0 s^0 - \mu^0 e\| \le \gamma_2\rho_0.$$

If $\|X_0s^0-\mu^0e\|$ is small enough, then we can choose these constants because $\sigma_1<1$. We define a neighborhood of the path of centers:

$$\begin{aligned} \mathcal{N} \;=\; & \{(x,y,s,\theta) \;:\; x>0,\; s>0,\; \theta>0, \\ & Ax=b+\theta(\bar{b}+\tilde{b}),\;\; A^Ty+s=c+\theta(\bar{c}+\tilde{c}),\;\; \left\|(\tilde{b},\;\tilde{c})\right\|_A \le \gamma_3\rho_0, \\ & \|Xs-\theta\mu^0e\| \le \gamma_2\theta\rho_0,\;\; \gamma_0\theta\mu^0 e\le Xs\le\gamma_1\theta\mu^0 e\}. \end{aligned}$$

We define $\theta^0=1$. It is easy to see that $(x^0,y^0,s^0,\theta^0)\in\mathcal{N}$ and $\mathcal{P}\subset\mathcal{N}$.

We briefly explain the various quantities in the above definition. The quantity γ_2 is a tolerance for the violation of the third equation in (3). By choosing $x^0=s^0=\sqrt{\mu^0}e$ as starting point we could choose $\gamma_2>0$ arbitrarily small resulting in a rather narrow neighborhood $\mathcal{N}$. If σ_1, σ_2 are small, γ_2 may be chosen large, and then two additional numbers γ_0 and γ_1 are needed to control the ∞-norm of the violation of the third equation in (3). Note that for $\gamma_2\rho_0<\mu^0$ we can remove the condition $\gamma_0\theta\mu^0e\le Xs\le\gamma_1\theta\mu^0e$ from the definition of $\mathcal{N}$. The number γ_3 controls the violation of the linear constraints, and γ_4 controls the ratio of "affine scaling direction" and "centering direction" in our algorithm below; the smaller γ_4, the larger the component of the affine scaling direction ($\theta'=0$) may be.

Our algorithm generates a sequence $\{(x^k,y^k,s^k,\theta^k)\}$ in the neighborhood $\mathcal{N}$ starting from the initial point (x^0,y^0,s^0,θ^0). Suppose that we have the k-th iterate $(x^k,y^k,s^k,\theta^k)\in\mathcal{N}$. We shall show how to compute the next iterate $(x^{k+1},y^{k+1},s^{k+1},\theta^{k+1})\in\mathcal{N}$. At the point (x^k,y^k,s^k), we would like to obtain a center which is a solution of (3) for a $\theta\in[0,\theta^k]$. So we try to compute the Newton direction $(\Delta x,\Delta y,\Delta s)$ for the system (3), that is, the solution of

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ S_k & 0 & X_k \end{pmatrix}\begin{pmatrix} \Delta x \\ \Delta y \\ \Delta s \end{pmatrix} = -\begin{pmatrix} Ax^k-b-\theta\bar{b} \\ A^Ty^k+s^k-c-\theta\bar{c} \\ X_ks^k-\theta\mu^0e \end{pmatrix}. \tag{4}$$

In order to compute a good approximation to $(x,y,s,\theta)\in\mathcal{P}$, we require to choose a value of the parameter θ such that the size of the right hand side of (4) is not too big. So we set

$$\theta' = \min\left\{\theta\in[\gamma_4\theta^k,\theta^k] \;:\; \sigma_1\left\|\begin{pmatrix} Ax^k-b-\theta\bar{b} \\ s^k-c-\theta\bar{c}\end{pmatrix}\right\|_A + \sigma_2\|X_ks^k-\theta\mu^0e\| \le \gamma_3\theta\rho_0\right\}. \tag{5}$$

From Assumption 1, we can compute an approximate solution $(\Delta x', \Delta y', \Delta s')$ of (4) for $\theta = \theta'$ such that

$$\left\| \begin{pmatrix} A(x^k + \Delta x') - b - \theta' \bar{b} \\ s^k + \Delta s' - c - \theta' \bar{c} \end{pmatrix} \right\|_{\mathcal{A}} \leq \sigma_1 \left\| \begin{pmatrix} Ax^k - b - \theta' \bar{b} \\ s^k - c - \theta' \bar{c} \end{pmatrix} \right\|_{\mathcal{A}} + \sigma_2 \| X_k s^k - \theta' \mu^0 e \|$$

and

$$S^k \Delta x' + X^k \Delta s' = -(X_k s^k - \theta' \mu^0 e). \tag{6}$$

Using the definition of θ', the above inequality implies

$$\left\| \begin{pmatrix} A(x^k + \Delta x') - b - \theta' \bar{b} \\ s^k + \Delta s' - c - \theta' \bar{c} \end{pmatrix} \right\|_{\mathcal{A}} \leq \gamma_3 \theta' \rho_0. \tag{7}$$

We choose a step size

$$\alpha^k = \max\{\alpha' \in [0,1] \ : \ (x^k, y^k, s^k, \theta^k) + \alpha(\Delta x', \Delta y', \Delta s', \theta' - \theta^k) \in \mathcal{N} \ \text{ for each } \ \alpha \in [0, \alpha']\}.$$

Then set

$$(x^{k+1}, y^{k+1}, s^{k+1}, \theta^{k+1}) = (x^k, y^k, s^k, \theta^k) + \alpha^k (\Delta x', \Delta y', \Delta s', \theta' - \theta^k).$$

Our algorithm is summarized as follows.

Algorithm Let (x^0, y^0, s^0) be the initial point. Set $k = 0$, $\mu^0 = (x^0)^T s^0 / n$, and $\theta^0 = 1$.

Step 1: Compute θ' by (5). Compute an approximate solution $(\Delta x', \Delta y', \Delta s')$ which satisfies (6) and (7).

Step 2: Compute a step size α^k and the next iterate $(x^{k+1}, y^{k+1}, s^{k+1}, \theta^{k+1})$ as shown above.

Step 3: Increase k by 1 and go to Step 1.

We point out that we need to specify the quantities $\gamma_0, \ldots \gamma_4$ and σ_1, σ_2 before the first iteration of the algorithm. In particular, we need an advance bound on σ_2.

The next theorem summarizes our main result, namely global linear convergence of our algorithm.

Theorem 1 *Let $\{(x^k, y^k, s^k, \theta^k)\}$ be a sequence generated by our algorithm. The sequence is bounded if and only if the primal-dual linear programming problem is feasible. If the sequence is bounded, then $\theta^k \to 0$ linearly as $k \to \infty$ and any accumulation point of $\{(x^k, y^k, s^k)\}$ is a solution of the problem.*

In this theorem, we do not assume how small the constants σ_1 and σ_2 introduced in Assumption 1 are. So we can solve the primal-dual linear programming problem under very rough inexact computation of the approximate solution $(\Delta x', \Delta y', \Delta s')$, which satisfies Assumption 1 for $\sigma_1 = .99$ and $\sigma_2 = 100$ for example. We will also show that the norms $\|U_k^{-1}\|$ defined in Section 2 are uniformly bounded, so that the substitution (2) is "compatible" with Assumption 1.

4 Discussion of the Main Result

The conceptual algorithm of the previous section and the main result do not address two important issues.

1. The main result depends on Assumption 1 which seems to be a very mild assumption at the first glance: In a certain norm associated with the linear system to be solved, this assumption requires a reduction of the residual by merely 1 % for example (when $\sigma_1 = 0.99$). Nevertheless we would like to know how difficult is it to satisfy this assumption.

2. The algorithm makes use of certain values θ' in (5) and α^k. How can we compute these values?

We will give a partial answer to both questions. A complete answer certainly depends on the type of computations (complete factorization versus iterative linear systems solver) that is used in the overall interior-point method.

For our discussion and our further analysis we factor the matrix A in (1). Let $A^T = Q_1 R$, where Q_1 is an $n \times m$ submatrix of an orthonormal $n \times n$ matrix $Q = (Q_1, Q_2)$ and R is a nonsingular $m \times m$ matrix. It follows from this definition that

$$\begin{aligned}
&Q_1 Q_1^T + Q_2 Q_2^T \ = \ I, \\
&\begin{pmatrix} Q_1^T Q_1 & Q_1^T Q_2 \\ Q_2^T Q_1 & Q_2^T Q_2 \end{pmatrix} \ = \ I, \\
&Q_2^T A^T \ = \ 0, \\
&R^{-1} Q_1^T A^T = I, \\
&\|Q_1\| \ \le \ 1, \quad \|Q_2\| \ \le \ 1,
\end{aligned}$$

We will use these relations throughout this paper without citation. System (1) is equivalent to

$$\begin{pmatrix} Q_1^T & 0 \\ 0 & Q_2^T \\ S_k & X_k \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta s \end{pmatrix} = \begin{pmatrix} R^{-T} p \\ Q_2^T q \\ r \end{pmatrix} \tag{8}$$

with $\Delta y = R^{-1} Q_1^T (q - \Delta s)$.

With these definitions, one readily verifies that

$$\left\|(\tilde{b}, \tilde{c})\right\|_A = \left\|(R^{-T}\tilde{b}, \ Q_2^T \tilde{c})\right\| \quad \text{for any} \ \ (\tilde{b}, \tilde{c})$$

and thus, the inequality in Assumption 1 is equivalent to

$$\|(Q_1^T \Delta x' - R^{-T} p, \ Q_2^T \Delta s' - Q_2^T q)\| \le \sigma_1 \|(R^{-T} p, \ Q_2^T q)\| + \sigma_2 \|r\|.$$

This formulation of Assumption 1 will be used in the analysis in Section 5.

To further relate Assumption 1 to standard error measures based on the norm of the residual, we substitute (2) into the above relation. To this end let the vectors x and s be partitioned as $x = (x^{(1)}, x^{(2)})$ and $s = (s^{(1)}, s^{(2)})$ such that $x^{(1)} > s^{(1)}$ componentwise and $x^{(2)} \le s^{(2)}$ componentwise. Likewise we partition the rows of the matrices Q_1 and Q_2. Define the $n \times n$ diagonal matrices

$$D_1 = \begin{pmatrix} -(X^{(1)})^{-1} S^{(1)} & 0 \\ 0 & I \end{pmatrix} \quad \text{and} \quad D_2 = \begin{pmatrix} I & 0 \\ 0 & -(S^{(2)})^{-1} X^{(2)} \end{pmatrix}.$$

Assumption 1 is then equivalent to

$$\left\| \begin{pmatrix} Q_1^T D_2 \\ Q_2^T D_1 \end{pmatrix} \begin{pmatrix} (\Delta x'')^{(1)} \\ (\Delta s'')^{(2)} \end{pmatrix} - \begin{pmatrix} R^{-T} p - (Q_1^{(2)})^T (S^{(2)})^{-1} r^{(2)} \\ Q_2^T q - (Q_2^{(1)})^T (X^{(1)})^{-1} r^{(1)} \end{pmatrix} \right\| \le \sigma_1 \left\| \begin{pmatrix} R^{-T} p \\ Q_2^T q \end{pmatrix} \right\| + \sigma_2 \|r\|.$$

We show in Lemma 9 below that $(X^{(1)})^{-1}$ and $(S^{(2)})^{-1}$ are uniformly bounded for all k. Since $Q = (Q_1, Q_2)$ is orthonormal, it follows for $\sigma_2 \ge \sup\{\left\|(X^{(1)})^{-1}\right\| + \left\|(S^{(2)})^{-1}\right\|\}\sigma_1$, that the above condition is slightly weaker than requiring

$$\|Bz - d\| \le \sigma_1 \|d\| \quad \text{where} \ \ B = \begin{pmatrix} Q_1^T D_1 \\ Q_2^T D_2 \end{pmatrix} \quad \text{and} \quad z = \begin{pmatrix} (\Delta x'')^{(1)} \\ (\Delta s'')^{(2)} \end{pmatrix}. \tag{9}$$

(Straightforward). This condition is stated in a standard form used for the error analysis for systems of linear equations. In particular, it is well-known that Gaussian elimination with partial pivoting is backward stable. Hence, if, for example, the linear system $Bz = d$ was solved by Gaussian elimination with partial pivoting, one could guarantee (9), and thus also Assumption 1 with σ_1 in the order of the machine precision, and a rather large value for σ_2. On the other hand, if the linear systems defining the search direction are solved by some iterative method, the condition number of B has a strong effect on the rate of convergence. Unfortunately, the condition number of the matrix B may be unbounded if the linear program is degenerate. Thus, in the final stage of our interior-point method when applied to solve a degenerate linear program, iterative linear systems solvers may use a high number of iterations in order to satisfy Assumption 1. Nevertheless, our analysis allows an interesting observation regarding the stability of interior-point methods:

1. The feature that one can solve a given linear program to a very high accuracy even when solving the linear systems at each step only to a very low accuracy (in the above residual norm (9)), this feature is not shared by the simplex method.
2. There are many discussions about ill-conditioning in interior-point methods. The above considerations show that this version of an interior-point method is quite insensitive to ill-conditioning: When the linear systems are ill-conditioned (in the final stage of the algorithm) the error, i.e. the difference of the approximate solution and the true solution of the linear system may be large, even if the residual happened to be small. Here, we make no assumption about the error, we do not even require that the residual is small (a reduction by just 1% is sufficient for our analysis), but nevertheless we can guarantee overall convergence.

In this paper, we do not specify how to compute the approximate solution $(\Delta x', \Delta y', \Delta s')$ which satisfies Assumption 1. We point out, however, that the matrix A is usually sparse, but the matrix Q may not be. Thus, while the matrix Q is a very suitable tool for our analysis below, we should not use it when we solve the system (1).

The above addresses the first issue about how difficult it may typically be to satisfy Assumption 1.

The second issue concerns how to determine θ' and α^k. If we are able to evaluate $\|(p,q)\|_A$ for a given pair of vectors, p and q, then both quantities can be approximated, for example, by some Armijo-type search. Also, the verification, that a given approximate solution $\Delta x', \Delta s'$ satisfies Assumption 1 can be reduced to the evaluation of $\|(p,q)\|_A$ for certain p and q.

If the interior-pint method is based on direct solvers for the linear equations, the evaluation of $\|(p,q)\|_A$ is fairly straightforward. The matrix R for example can also be obtained from a Cholesky decomposition of AA^T and is available in some implementations of interior-point methods. Given R, one backsolve returns $R^{-T}p$, and two further backsolves for $R^TRy = Aq$ yield the minimizing vector y in the definition of $\|\,.\,\|_A$, so that $\|(p,q)\|_A$ is available with at most three backsolves.

If the interior-point method is based on iterative linear systems solvers, upper bounds for $\left\|Q_2^T(\Delta s' - q)\right\|$ can be obtained from conjugate gradient methods for solving $AA^Ty = Aq$, while cg-methods for the (singular semidefinite) system $A^TAz = A^Tp$ yield lower bounds for $\left\|R^{-T}(A\Delta x' - p)\right\|$. We point out that both linear systems above have a constant condition number independent of the iteration index k. By preprocessing the matrix A prior to solving the linear program one may further control the singular values of A. (An extreme form of preprocessing would consist of premultiplying the linear equation $Ax = b$ by R^{-T}, resulting in the equivalent relation $Q_1^Tx = R^{-T}b$ with all singular values of Q_1^T equal to 1. As mentioned before this form of preprocessing is not practical when A is sparse.) We do not discuss these details any further but continue with proving convergence of the algorithm.

5 Global Convergence

In this section, as our main result, we prove global linear convergence of our algorithm. In the final part of this section we also show that the norms $\|U_k^{-1}\|$ defined in Section 2 are uniformly bounded.

We prove Theorem 1 by the following five steps:

(i) If the problem is feasible, the sequence $\{(x^k, y^k, s^k, \theta^k)\}$ is bounded.

(ii) If there exist constants $\tau \in (0,1)$ and $\hat{\alpha} \in (0,1)$ independent of k such that $\theta' \le \tau\theta^k$ and $\alpha^k \ge \hat{\alpha}$ for each k, then $\theta^k \to 0$ linearly as $k \to \infty$, and any accumulation point of the sequence $\{(x^k, y^k, s^k)\}$ is a solution of the problem.

(iii) There exists a constant $\tau \in (0,1)$ such that $\theta' \le \tau\theta^k$.

(iv) If the sequence $\{\|\Delta X'\Delta s'\|/\theta^k\}$ is bounded, there exists a constant $\hat{\alpha} \in (0,1)$ such that $\alpha^k \ge \hat{\alpha}$.

(v) If $\{(x^k, y^k, s^k)\}$ is bounded, then $\{\|\Delta X'\Delta s'\|/\theta^k\}$ is also bounded.

Since the primal-dual linear programming has a solution if it is feasible, **(i)** follows from the next two lemmas.

Lemma 1 *For any $\tilde{b} \in R^m$ and $\tilde{c} \in R^n$ such that $\left\|(\tilde{b}, \tilde{c})\right\|_A \le \gamma_3\rho_0$, there exists $(\tilde{x}, \tilde{y}, \tilde{s})$ which satisfies*

$$\begin{gathered} A\tilde{x} = b + \bar{b} + \tilde{b}, \\ A^T\tilde{y} + \tilde{s} = c + \bar{c} + \tilde{c}, \\ (1-\gamma_3)\rho_0 e \le \tilde{x} \le (1+\gamma_3)x^0, \\ (1-\gamma_3)\rho_0 e \le \tilde{s} \le (1+\gamma_3)s^0. \end{gathered}$$

Proof: This lemma easily follows form the definition of $\|\cdot\|_A$. In fact, all the conditions hold true for

$$(\tilde{x}, \tilde{y}, \tilde{s}) = (x^0 + Q_1 R^{-T}\tilde{b}, y^0 + R^{-1}Q_1^T\tilde{c}, s^0 + Q_2 Q_2^T\tilde{c}).$$

Q.E.D.

Lemma 2 *If the primal-dual linear programming problem has a solution (x^*, y^*, s^*) then*

$$\|(x,s)\|_1 \le \frac{1}{(1-\gamma_3)\rho_0}\left((1+\gamma_3)((s^0)^T x^* + (x^0)^T s^*) + \theta(1+\gamma_3)^2 n\mu^0 + \gamma_1 n\mu^0\right)$$

for each $\theta \in (0,1]$ and $(x, y, s, \theta) \in \mathcal{N}$.

Proof: We define

$$\begin{aligned} \tilde{b} &= (Ax-b)/\theta - \bar{b} \\ \tilde{c} &= (A^T y + s - c)/\theta - \bar{c}. \end{aligned}$$

Since $(x, y, s, \theta) \in \mathcal{N}$, we have that $\left\|(\tilde{b}, \tilde{c})\right\|_A \le \gamma_3\rho_0$. So there exists $(\tilde{x}, \tilde{y}, \tilde{s})$ which satisfies the conditions in Lemma 1. Then

$$A((1-\theta)x^* + \theta\tilde{x} - x) = 0$$

and

$$A^T((1-\theta)y^* + \theta\tilde{y} - y) + ((1-\theta)s^* + \theta\tilde{s} - s) = 0.$$

Hence we have that

$$((1-\theta)x^* + \theta\tilde{x} - x)^T((1-\theta)s^* + \theta\tilde{s} - s) = 0$$

or equivalently

$$((1-\theta)x^* + \theta\tilde{x})^T s + ((1-\theta)s^* + \theta\tilde{s})^T x = ((1-\theta)x^* + \theta\tilde{x})^T((1-\theta)s^* + \theta\tilde{s}) + x^T s. \tag{10}$$

Using $(x^*)^T s^* = 0$, $(1-\gamma_3)\rho_0 e \le \tilde{x} \le (1+\gamma_3)x^0$, $(1-\gamma_3)\rho_0 e \le \tilde{s} \le (1+\gamma_3)s^0$, and $Xs \le \gamma_1\theta\mu^0 e$, we have that

$$\begin{aligned} \theta(1-\gamma_3)\rho_0(e^T s + e^T x) &\le ((1-\theta)x^* + \theta\tilde{x})^T s + ((1-\theta)s^* + \theta\tilde{s})^T x \\ &= ((1-\theta)x^* + \theta\tilde{x})^T((1-\theta)s^* + \theta\tilde{s}) + x^T s \\ &\le \theta(1-\theta)(1+\gamma_3)((s^0)^T x^* + (x^0)^T s^*) + \theta^2(1+\gamma_3)^2(x^0)^T s^0 + \gamma_1\theta n\mu^0. \end{aligned}$$

The inequality in the lemma follows from this inequality and $\theta \in (0,1]$. **Q.E.D.**

Suppose that there exist constants $\tau \in (0,1)$ and $\hat{\alpha} \in (0,1)$ independent of k such that $\theta' \leq \tau\theta^k$ and $\alpha^k \geq \hat{\alpha}$ for each k. Then we have that

$$\begin{aligned} \theta^{k+1} &= \theta^k + \alpha^k(\theta' - \theta^k) \\ &\leq \theta^k + \hat{\alpha}(\tau - 1)\theta^k \\ &\leq (1 - \hat{\alpha}(1-\tau))\theta^k. \end{aligned} \tag{11}$$

Since $1 - \hat{\alpha}(1-\tau) < 1$, we have that $\theta^k \to 0$ linearly as $k \to \infty$. Then $\|X_k s^k\| \to 0$, $\|Ax^k - b\| \to 0$, $\|A^T y^k + s^k - c\| \to 0$ as $k \to \infty$ because $(x^k, y^k, s^k, \theta^k) \in \mathcal{N}$. Hence we have shown **(ii)**.

The statement **(iii)** follows from the next lemma.

Lemma 3 *Define*

$$\lambda = \sigma_1 \left\|(\bar{b},\ \bar{c})\right\|_{\mathcal{A}} + \sigma_2\sqrt{n}\mu^0$$

and

$$\tau_1 = \frac{(\sigma_1\gamma_3 + \sigma_2\gamma_2)\rho_0 + \lambda}{\gamma_3\rho_0 + \lambda}.$$

Then $\tau_1 < 1$, *and* $\theta' \leq \max\{\gamma_4, \tau_1\}\theta^k$ *at each iteration of the algorithm.*

Proof: Since $\sigma_1\gamma_3 + \sigma_2\gamma_2 < \gamma_3$, we have that $\tau_1 < 1$. For each $\theta \in [\tau_1\theta^k, \theta^k]$, we have from the definition of $\mathcal{N}$ that

$$\begin{aligned} &\sigma_1 \left\|(Ax^k - b - \theta\bar{b},\ s^k - c - \theta\bar{c})\right\|_{\mathcal{A}} + \sigma_2\|X_k s^k - \theta\mu^0 e\| - \gamma_3\theta\rho_0 \\ &\leq \sigma_1 \left\|(Ax^k - b - \theta^k\bar{b},\ s^k - c - \theta^k\bar{c})\right\|_{\mathcal{A}} + \sigma_2\|X_k s^k - \theta^k\mu^0 e\| \\ &\quad + (\theta^k - \theta)(\sigma_1 \left\|(\bar{b},\ \bar{c})\right\|_{\mathcal{A}} + \sigma_2\|\mu^0 e\|) - \gamma_3\theta\rho_0 \\ &\leq \theta^k\sigma_1\gamma_3\rho_0 + \theta^k\sigma_2\gamma_2\rho_0 + (\theta^k - \theta)\lambda - \gamma_3\theta\rho_0 \\ &\leq 0. \end{aligned}$$

Hence $\theta' \leq \max\{\gamma_4, \tau_1\}\theta^k$ by the definition. **Q.E.D.**

The statement **(iv)** follows from the next two lemmas and $\theta' \geq \gamma_4\theta^k$.

Lemma 4 *For each* $\alpha \in [0,1]$, *we define*

$$\begin{aligned} \tilde{b}(\alpha) &= \frac{A(x^k + \alpha\Delta x') - b}{(1-\alpha)\theta^k + \alpha\theta'} - \bar{b}, \\ \tilde{c}(\alpha) &= \frac{A^T(y^k + \alpha\Delta y') + (s^k + \alpha\Delta s') - c}{(1-\alpha)\theta^k + \alpha\theta'} - \bar{c}. \end{aligned}$$

Then

$$\left\|(\tilde{b}(\alpha),\ \tilde{c}(\alpha))\right\|_{\mathcal{A}} \leq \gamma_3\rho_0.$$

Proof: Using $(x^k, y^k, s^k, \theta^k) \in \mathcal{N}$ and (7), we see that

$$\begin{aligned} &((1-\alpha)\theta^k + \alpha\theta') \left\|(\tilde{b}(\alpha),\ \tilde{c}(\alpha))\right\|_{\mathcal{A}} \\ &= \left\|(A(x^k + \alpha\Delta x') - b - ((1-\alpha)\theta^k + \alpha\theta')\bar{b},\ s^k + \alpha\Delta s' - c - ((1-\alpha)\theta^k + \alpha\theta')\bar{c})\right\|_{\mathcal{A}} \\ &\leq (1-\alpha) \left\|(Ax^k - b - \theta^k\bar{b},\ s^k - c - \theta^k\bar{c})\right\|_{\mathcal{A}} \\ &\quad + \alpha \left\|(A(x^k + \Delta x') - b - \theta'\bar{b},\ s^k + \Delta s' - c - \theta'\bar{c})\right\|_{\mathcal{A}} \\ &\leq (1-\alpha)\gamma_3\theta^k\rho_0 + \alpha\gamma_3\theta'\rho_0 \\ &= ((1-\alpha)\theta^k + \alpha\theta')\gamma_3\rho_0. \end{aligned}$$

Q.E.D.

Lemma 5 *At the k-th iteration of the algorithm, we define*

$$\hat{\alpha}^k = \min\left\{ \frac{\theta'\gamma_2\rho_0}{\|\Delta X'\Delta s'\|}, \frac{\theta'(1-\gamma_0)\mu^0}{\|\Delta X'\Delta s'\|_\infty}, \frac{\theta'(\gamma_1-1)\mu^0}{\|\Delta X'\Delta s'\|_\infty} \right\}.$$

Then $\alpha^k \geq \hat{\alpha}^k$.

Proof: Suppose that $\alpha \in [0, \hat{\alpha}^k]$. Using (6), we see that

$$\begin{aligned}
&\|(X_k + \alpha\Delta X')(s^k + \alpha\Delta s') - (\theta^k + \alpha(\theta' - \theta^k))\mu^0 e\| \\
&\quad = \|X_k s^k + \alpha(\theta'\mu^0 e - X_k s^k) + \alpha^2 \Delta X'\Delta s' - (\theta^k + \alpha(\theta' - \theta^k))\mu^0 e\| \\
&\quad \leq (1-\alpha)\|X_k s^k - \theta^k \mu^0 e\| + \alpha^2\|\Delta X'\Delta s'\| \\
&\quad \leq (1-\alpha)\gamma_2\theta^k\rho_0 + \alpha\theta'\gamma_2\rho_0 \\
&\quad = \gamma_2(\theta^k + \alpha(\theta' - \theta^k))\rho_0, \\
&(X_k + \alpha\Delta X')(s^k + \alpha\Delta s') - (\theta^k + \alpha(\theta' - \theta^k))\gamma_0\mu^0 e \\
&\quad = X_k s^k + \alpha(\theta'\mu^0 e - X_k s^k) + \alpha^2\Delta X'\Delta s' - (\theta^k + \alpha(\theta' - \theta^k))\gamma_0\mu^0 e \\
&\quad = (1-\alpha)(X_k s^k - \theta^k\gamma_0\mu^0 e) + \alpha\theta'(1-\gamma_0)\mu^0 e + \alpha^2\Delta X'\Delta s' \\
&\quad \geq 0, \\
&(X_k + \alpha\Delta X')(s^k + \alpha\Delta s') - (\theta^k + \alpha(\theta' - \theta^k))\gamma_1\mu^0 e \\
&\quad = (1-\alpha)(X_k s^k - \theta^k\gamma_1\mu^0 e) - \alpha\theta'(\gamma_1 - 1)\mu^0 e + \alpha^2\Delta X'\Delta s' \\
&\quad \leq 0.
\end{aligned}$$

The second relation also implies that $(x^k + \alpha\Delta x', s^k + \alpha\Delta s') > 0$ from the continuity with respect to α. Combining these results and Lemma 4, we have that $(x^k, y^k, s^k, \theta^k) + \alpha(\Delta x', \Delta y', \Delta s', \theta' - \theta^k) \in \mathcal{N}$ for each $\alpha \in [0, \hat{\alpha}^k]$. Hence $\alpha^k \geq \hat{\alpha}^k$ by the definition. **Q.E.D.**

The statement **(v)** follows from the next three lemmas.

Lemma 6 *The solution of (1) is expressed as*

$$\begin{aligned}
D^{-1}\Delta x &= PD^{-1}Q_1R^{-T}p - (I-P)DQ_2Q_2^Tq + (I-P)(X_kS_k)^{-.5}r, \\
\Delta y &= R^{-1}Q_1^Tq + (AD^2A^T)^{-1}AD(D^{-1}Q_1R^{-T}p + DQ_2Q_2^Tq - (X_kS_k)^{-.5}r), \\
D\Delta s &= -PD^{-1}Q_1R^{-T}p + (I-P)DQ_2Q_2^Tq + P(X_kS_k)^{-.5}r,
\end{aligned}$$

where $D = X_k^{.5}S_k^{-.5}$ *and* $P = DA^T(AD^2A^T)^{-1}AD$.

Proof: Suppose that $(\Delta x, \Delta y, \Delta s)$ is expressed as above. Since $ADP = AD$, $AD(I-P) = 0$, and $A = R^TQ_1^T$, we see that

$$\begin{aligned}
A\Delta x &= AD(D^{-1}\Delta x) \\
&= ADD^{-1}Q_1R^{-T}p \\
&= p, \\
A^T\Delta y + \Delta s &= D^{-1}(DA^T\Delta y + D\Delta s) \\
&= D^{-1}(DQ_1Q_1^Tq + P(D^{-1}Q_1R^{-T}p + DQ_2Q_2^Tq - (X_kS_k)^{-.5}r) + D\Delta s) \\
&= D^{-1}(DQ_1Q_1^Tq + DQ_2Q_2^Tq) \\
&= q, \\
S_k\Delta x + X_k\Delta s &= (X_kS_k)^{.5}(D^{-1}\Delta x + D\Delta s) \\
&= (X_kS_k)^{.5}(X_kS_k)^{-.5}r \\
&= r.
\end{aligned}$$

Q.E.D.

Lemma 7 *At the k-th iteration of the algorithm, let*

$$\begin{aligned}
\tilde{b}^k &= (Ax^k - b)/\theta^k - \bar{b},\\
\tilde{c}^k &= (A^T y^k + s^k - c)/\theta^k - \bar{c},\\
\hat{b} &= (A(x^k + \Delta x') - b)/\theta' - \bar{b},\\
\hat{c} &= (A^T(y^k + \Delta y') + (s^k + \Delta s') - c)/\theta' - \bar{c}.
\end{aligned}$$

Then $\|(\tilde{b}^k,\ \tilde{c}^k)\|_A \le \gamma_3\rho_0$ *and* $\|(\hat{b},\ \hat{c})\|_A \le \gamma_3\rho_0$, *and* $(\Delta x', \Delta y', \Delta s')$ *is the solution of System (1) for*

$$\begin{aligned}
p &= (\theta' - \theta^k)\bar{b} + \theta'\hat{b} - \theta^k\tilde{b}^k,\\
q &= (\theta' - \theta^k)\bar{c} + \theta'\hat{c} - \theta^k\tilde{c}^k,\\
r &= -(X_k s^k - \theta'\mu^0 e).
\end{aligned}$$

Proof: It is straightforward using Lemma 4. **Q.E.D.**

Lemma 8 *If* $\{(x^k, y^k, s^k)\}$ *is bounded, then* $\{\|\Delta X'\Delta s'\|/\theta^k\}$ *is also bounded.*

Proof: Suppose that $\{(x^k, y^k, s^k)\}$ is bounded. Then a constant $\eta_1 > 0$ exists such that

$$x_i^k \le \eta_1 \text{ and } s_i^k \le \eta_1 \text{ for each } i \text{ and } k. \tag{12}$$

Since $X_k s^k \ge \gamma_0\theta^k\mu^0 e$, we have that

$$\xi_1\theta^k \le x_i^k \text{ and } \xi_1\theta^k \le s_i^k \text{ for each } i \text{ and } k,$$

where $\xi_1 = \gamma_0\mu^0/\eta_1$. Let $d_i = \sqrt{x_i^k/s_i^k}$ be the i-th diagonal component of D at the k-th iteration of the algorithm. Then

$$\sqrt{\theta^k}/\eta_2 \le d_i \le \eta_2/\sqrt{\theta^k},$$

where $\eta_2 = \sqrt{\eta_1/\xi_1}$. Hence

$$\|D\| \le \eta_2/\sqrt{\theta^k} \text{ and } \|D^{-1}\| \le \eta_2/\sqrt{\theta^k}.$$

We define (p, q, r) as in Lemma 7. Then a constant $\eta_3 > 0$ exists such that

$$\|(R^{-T}p,\ Q_2^T q)\| = \|(p, q)\|_A \le \eta_3\theta^k,\ \|r\| \le \eta_3\theta^k.$$

Since $(\Delta x', \Delta y', \Delta s')$ is the solution of (1) for the (p, q, r), from Lemma 6 we see that

$$\begin{aligned}
\|D^{-1}\Delta x'\| &\le \|D^{-1}\|\|R^{-T}p\| + \|D\|\|Q_2^T q\| + \|(X_k S^k)^{-.5}\|\|r\|\\
&\le \frac{\eta_2}{\sqrt{\theta^k}}\eta_3\theta^k + \frac{\eta_2}{\sqrt{\theta^k}}\eta_3\theta^k + \frac{1}{\sqrt{\gamma_0\theta^k\mu^0}}\eta_3\theta^k\\
&= \eta_4\sqrt{\theta^k}
\end{aligned}$$

for $\eta_4 = (2\eta_2\eta_3 + \eta_3/\sqrt{\gamma_0\mu^0})$, where we have used the relations $\|P\| \le 1$, $\|I - P\| \le 1$, $\|Q_1\| \le 1$, and $\|Q_2\| \le 1$. Similarly we have the same bound of $\|D\Delta s'\|$. Hence

$$\|\Delta X'\Delta s'\| \le \|D^{-1}\Delta x'\|\|D\Delta s'\| \le \eta_4^2\theta^k.$$

Q.E.D.

To close this section, we prove that $\|U_k^{-1}\|$ defined in Section 2 is bounded. It follows from the next lemma.

Lemma 9 *If the primal-dual linear programming problem is feasible,* $\zeta > 0$ *exists such that*

$$\max\{x_i, s_i\} \ge \zeta$$

for any $\theta \in (0, 1]$, $(x, y, s, \theta) \in \mathcal{N}$ *and* i.

Proof: It is well-known that if the primal-dual linear programming problem is feasible then a strictly complementarity solution (x^*, y^*, s^*) of the problem exists, i.e., $\zeta_1 > 0$ exists such that $\max\{x_i^*, s_i^*\} \geq \zeta_1$ for each i. Define $\zeta_2 = \min\{\zeta_1, \rho_0(1-\gamma_3)\}$ and let $(\tilde{x}, \tilde{y}, \tilde{s})$ be defined as in the proof of Lemma 2. From (10) and Lemma 1, we see that

$$\begin{aligned} \zeta_2 \min\{x_i, s_i\} &\leq ((1-\theta)x^* + \theta\tilde{x})^T s + ((1-\theta)s^* + \theta\tilde{s})^T x \\ &= ((1-\theta)x^* + \theta\tilde{x})^T((1-\theta)s^* + \theta\tilde{s}) + x^T s \\ &\leq \theta(1-\theta)(\tilde{s}^T x^* + \tilde{x}^T s^*) + \theta^2 \tilde{x}^T \tilde{s} + \gamma_1 \theta n \mu^0 \\ &\leq \theta\eta \end{aligned}$$

for a constant $\eta > 0$ independent of the point (x, y, s, θ). Since $Xs \geq \gamma_0 \theta \mu^0 e$, we obtain that

$$\max\{x_i, s_i\} \geq \gamma_0 \zeta_2 \mu^0 / \eta.$$

Q.E.D.

6 Polynomial-Time Convergence

In this section, we prove the following convergence theorem. We use the notations $g_1 = O(f(n))$, $g_2 = \Omega(f(n))$, and $g_3 = \Theta(f(n))$ for a function f of n, which imply that positive constants ω_0 and ω_1 exist such that

$$g_1 \leq \omega_1 f(n), \ g_2 \geq \omega_0 f(n), \ \omega_0 f(n) \leq g_3 \leq \omega_1 f(n).$$

Theorem 2 *Let $f(n)$ be a function of n, and let $\epsilon > 0$ be a small constant. Let some linear program with n unknowns be given. Suppose that we use an initial point (x^0, y^0, s^0) and $\gamma_2 > 0$ such that*

$$\gamma_2 = \Omega(\rho_0) \ \textit{and} \ \ \|(x^0, s^0)\|_\infty = \Theta(\rho_0), \tag{13}$$

and that the constants γ_0, γ_1, γ_3, and γ_4 are independent of the data. If $\delta \in (0, \gamma_3)$ is independent of the data and $\sigma_1 \in [0, 1)$ and $\sigma_2 \geq 0$ are small enough such that

$$\sigma_1\gamma_3 + \sigma_2\gamma_2 + \delta \leq \gamma_3 \quad \textit{and} \quad \lambda = \sigma_1 \left\|(\bar{b}, \bar{c})\right\|_A + \sigma_2 \sqrt{n}\mu^0 \leq f(n)\rho_0,$$

and if there exists a solution (x^*, y^*, s^*) of the primal-dual linear programming problem such that

$$\|(x^*, s^*)\|_\infty = O(\rho_0), \tag{14}$$

then $\theta^k \leq \epsilon$ for $k = O\left(n^2(1 + f(n))\ln(1/\epsilon)\right)$.

From this theorem, if we can compute the exact solution of the linear system of equations, i.e., $\sigma_1 = 0$ and $\sigma_2 = 0$, then the number of iterations of our algorithm is $O(n^2 \ln(1/\epsilon))$ to get an ϵ-approximate solution. This bound of our algorithm is equal to the one of the infeasible-interior-point algorithms proposed by Zhang [8] and Mizuno [5]. This theorem extends the results in [8, 5] since it gives a sufficient condition to achieve the bound of $O(n^2 \ln(1/\epsilon))$ iterations under inexact computations. If we can compute an approximate solution, which satisfies the conditions in this theorem for $f(n) = constant$, at each iteration, then the number of iterations is bounded by $O(n^2 \ln(1/\epsilon))$.

Proof: From (13) and (14), we have that

$$\mu^0 = \Theta(\rho_0^2) \text{ and } (s^0)^T x^* + (x^0)^T s^* = O(n\rho_0^2).$$

From this relation and Lemma 2, we have that

$$\begin{aligned} \|(x^k, s^k)\|_1 &\leq \frac{1}{(1-\gamma_3)\rho_0}\left((1+\gamma_3)((s^0)^T x^* + (x^0)^T s^*) + \theta^k (1+\gamma_3)^2 n\mu^0 + \gamma_1 n \mu^0\right) \\ &= O(n\rho_0). \end{aligned}$$

So we see (12) for $\eta_1 = O(n\rho_0)$. By using the same discussion in the proof of Lemma 8, we have that

$$\|D\| \leq \eta_2/\sqrt{\theta^k} \text{ and } \|D^{-1}\| \leq \eta_2/\sqrt{\theta^k},$$

where $\eta_2 = O(\eta_1/\sqrt{\mu^0}) = O(n)$. We define (p, q, r) as in Lemma 7. Then

$$\begin{aligned}
\|PD^{-1}Q_1R^{-T}p\| &= \|PD^{-1}Q_1R^{-T}((\theta' - \theta^k)\bar{b} + \theta'\hat{b} - \theta^k\tilde{b}^k)\| \\
&\leq (\theta^k - \theta')\|DA^T(AD^2A^T)^{-1}AQ_1R^{-T}(Ax^0 - b)\| \\
&\quad +\theta'\|D^{-1}\|\|R^{-T}\hat{b}\| + \theta^k\|D^{-1}\|\|R^{-T}\tilde{b}^k\| \\
&\leq \theta^k\|DA^T(AD^2A^T)^{-1}(Ax^0 - Ax^*)\| + (\theta'/\sqrt{\theta^k})\eta_2\gamma_3\rho_0 + \sqrt{\theta^k}\eta_2\gamma_3\rho_0 \\
&\leq \theta^k\|PD^{-1}(x^0 - x^*)\| + 2\sqrt{\theta^k}\eta_2\gamma_3\rho_0 \\
&\leq \theta^k\|(X_kS_k)^{-.5}\|\|S_k(x^0 - x^*)\| + 2\sqrt{\theta^k}\eta_2\gamma_3\rho_0 \\
&\leq (\theta^k/\sqrt{\gamma_0\theta^k\mu^0})(\|x^0\|_\infty + \|x^*\|_\infty)\|s^k\|_1 + 2\sqrt{\theta^k}\eta_2\gamma_3\rho_0 \\
&= O(n\rho^0\sqrt{\theta^k}).
\end{aligned}$$

Similarly we can show that

$$\|(I - P)DQ_2Q_2^Tq\| = O(n\rho^0\sqrt{\theta^k}).$$

We also have that

$$\|(X_kS_k)^{-.5}r\| = O(\sqrt{n}\rho_0\sqrt{\theta^k})$$

So we see that

$$\|D^{-1}\Delta x'\| = O(n\rho_0\sqrt{\theta^k}).$$

Similarly we can obtain the same bound of $\|D\Delta s'\|$. Hence

$$\|\Delta X'\Delta s'\| = O(n^2\rho_0^2\theta^k),$$

which implies $1/\hat{\alpha} = O(n^2)$ from Lemma 5. From Lemma 3, we see that

$$1 - \tau_1 = \frac{(\gamma_3 - \sigma_1\gamma_3 - \sigma_2\gamma_2)}{\gamma_3 + \lambda/\rho_0},$$

which implies $1/(1 - \tau_1) = O(1 + f(n))$. Since

$$\theta^k \leq (1 - \hat{\alpha}(1 - \tau))^k$$

for $\tau = \max\{\gamma_4, \tau_1\}$ from Lemma 3 and (11), we have that $\theta^k \leq \epsilon$ for $k = (\hat{\alpha}(1 - \tau))^{-1}\ln(1/\epsilon) = O(n^2(1 + f(n))\ln(1/\epsilon))$. **Q.E.D.**

Acknowledgements

The second author would like to thank Roland Freund for many stimulating discussions.

References

[1] R.W. Freund, F. Jarre, and S. Mizuno, "Convergence of inexact interior-point algorithms for linear programs", in preparation, April 1996.

[2] N. Karmarkar, "A new polynomial-time algorithm for linear programming," *Combinatorica* 4 (1984) 373-395.

[3] M. Kojima, N. Megiddo and S. Mizuno, "A primal-dual infeasible-interior-point algorithm for linear programming," *Mathematical Programming* **61** (1993) 261-280.

[4] I. J. Lustig, R. E. Marsten and D. F. Shanno, "Computational experience with a primal-dual interior point method for linear programming," *Linear Algebra and Its Applications* **152** (1991) 191–222.

[5] S. Mizuno, "Polynomiality of infeasible-interior-point algorithms for linear programming," *Mathematical Programming* **67** (1994) 109-119.

[6] S. Mizuno, M. J. Todd, and Y. Ye, "A surface of analytic centers and infeasible-interior-point algorithms for linear programming," *Mathematics of Operations Research* **20** (1995) 52–67.

[7] K. Tanabe, "Centered Newton method for linear programming: Exterior point method (in Japanese)," *Proceedings of the Institute of Statistical Mathematics* **37** (1989) 146-148.

[8] Y. Zhang, "On the convergence of a class of infeasible interior-point methods for the horizontal linear complementarity problem," *SIAM Journal on Optimization* **4** (1994) 208–227.